

R

判別分析の例

■ テキストファイルから言語特徴量の算出

- ・ 学習者データの処理
 - ・ 作業ディレクトリーをデータの入っているディレクトリーに設定
 - ・ `list.files()` でファイルが入っていることを確認
- ・ 基礎的言語特徴量の算出プログラム (`myTextIndex.R`) の実行

```
myTextIndex()
```

- ・ 結果の保存ファイルを、一つ上のディレクトリーに作成する
- ・ 保存された結果のデータを読み込む
- ・ カラム名をつける
- ・ 同様に母語話者データも処理する

■ ファイル名の代わりに、カテゴリーを代入 (バックアップを作り作業)

■ 二種類のファイルを統合

■ VIF を計算してみる

- ・ Token と Type, GI が 10 以上
- ・ Token を残して、Type と GI を除く

判別分析

■ `lda()`

- ・ Leave-One-Out Cross Validation 付きで判別分析を行う

```
モデル <- lda( カテゴリー ~ ., data = データ )
```

■ 青木先生の `disc()`

判別分析 VIF を考慮した場合

■ `lda()`

■ `disc()`

- ・ プロットしてみる

Stepwise の判別分析

■ 青木先生の `sdis()`

```
source("http://aoki2.si.gunma-u.ac.jp/R/src/sdis.R", encoding="euc-jp")
```

■ グラフの描写